The Utility of a Realist Evaluation Approach in Implementing and Evaluating Health Equity Policy

Sanjeev Sridharan
St. Michael's Hospital and
University of Toronto
Toronto, Ontario

Tim Warren
Self Management Directorate,
Scottish Government
Edinburgh, Scotland

Abstract: *This article discusses the utility of realist evaluation in influencing policies aimed at reducing health inequities and improving population health. Using an example of a National Demonstration program in Scotland, the article discusses how realist evaluation can help in several aspects of policy implementation: exploring "loss in translation" from policy aspirations to program design, interrogating the program design, developing a range of learnings from conducting the evaluation, and aligning the learning with policy priorities. Conditions under which evaluations can lead to positive influence on policy makers and future policies are discussed.*

Keywords: *policy alignment, policy influence, positive evaluation, realist evaluation*

Résumé : *Cet article traite l'utilité de l'évaluation réaliste à influencer les politiques visant à réduire les inégalités sociales de la santé et à améliorer la santé de la population. En prenant l'exemple d'un programme national de démonstration en Écosse, cet article explique comment l'évaluation réaliste peut aider dans les aspects suivants de la mise en œuvre de la politique : à découvrir ce qui est « perdu dans la traduction » des aspirations politiques vers la conception des programmes; à questionner la conception du programme; à développer une gamme de leçons apprises de la conduite de l'évaluation; et à aligner l'apprentissage avec les priorités de la politique. L'article examine les conditions favorisant une influence positivedes évaluations sur les décideurs et les politiques de l'avenir.*

Mots clés : *alignement de la politique, influence de la politique, évaluation positive, évaluation réaliste*

Corresponding author: Sanjeev Sridharan, The Evaluation Centre for Complex Health Interventions, Keenan Research, Centre in the Li Ka Shing Knowledge Institute at St. Michael's Hospital and University of Toronto, ON, Canada; sridharans@smh.ca

 Canadian Journal of Program Evaluation / La Revue canadienne d'évaluation de programme
29.2 (fall / automne), 87–103 doi: 10.3138/cjpe.29.2.87

INTRODUCTION

The original call for submissions for this special theme segment of *CJPE* was to explore "how a realist evaluation approach (Pawson & Tilley, 1997) can support 'positive impacts' and also support 'positive thinking and action.'" The realist evaluation approach shifts focus from "Does a program work?" to "What is it about a program that makes it work?" A fundamental aspect of conducting realist evaluation (Pawson & Sridharan, 2009) is to explore the contexts, mechanisms, and outcomes that underlie programs.

Positive can mean different things to different people. One view of positive is that the experience of doing an evaluation is an empowering experience for the policy maker, program staff, and evaluators. A slightly different view of positive is that learning and possibly even program improvement happens as a result of the evaluation. As proof of such a positive impact of an evaluation, the learning and program improvement as a result of the evaluation continues well after the evaluation has ended.

We approached this invitation to think of positive approaches with some trepidation. It is rare that an evaluation approach is in itself positive or negative. It is often not the technical aspects of the approach but rather the relationships and trust that develop between the program planners and implementers, the commissioners of an evaluation, and the evaluation team that makes an evaluation approach positive. We viewed this article as an opportunity to reflect on the particulars of a realist evaluation that can result in program improvement.

The article also connects with several ideas presented by other articles in this special section—for example, MacCoy's focus on reframing as part of Appreciative Inquiry ("with a positive frame, people can open their minds to seeing new connections between ideas, people, and situations, often resulting in a flash of insight that is generative"), Stame's focus on emergent learning ("But positive thinking approaches to evaluation can do more than that; by challenging the "deficit oriented" theory implicit in many programs, they can provide tools for understanding how in the implementation of a program it is possible to expand the realm of the possible"), and Perrin's call to take a constructive approach to solving problems ("Problems and challenging findings are not neglected, but framed and approached from the perspective of what one can learn from what has taken place and what one can do about it in the future"). The purpose of this article is to reflect on the multiple rich ways in which thinking about contexts and mechanisms can help reach policy goals.

This article serves as an illustration of how a realist evaluation approach can be applied in practice. It discusses an evaluation focused on a health equity initiative in Paisley, Scotland, called *Have a Heart Paisley* (Sridharan et al., 2008a). We use *positive* in the sense of the evaluation generating learning that is used to improve program and/or policy on addressing health equities through programmatic interventions. In the spirit of a positive evaluation, this article is a collaboration between a policy lead and an evaluator. Sanjeev led the Evaluation team that conducted the evaluation of HaHP. Tim was the Policy Lead of Health Inequities in Scotland when the evaluation was completed.

The article starts by introducing the policy and programmatic aspects of Have a Heart Paisley (HaHP); we then briefly introduce realist evaluation and discuss why it can be useful in evaluating health inequities; we then discuss the utility of realist evaluation in policy implementation; we finally reflect on features of a realist evaluation that are necessary for an evaluation to be positive.

A BRIEF BACKGROUND TO *HAVE A HEART PAISLEY*

Policy Background

An empirical manifestation of inequalities in health is the inverse care law (Hart, 1971), which states that the availability of medical services is inversely related to the need in the population served. Mortality from coronary heart disease (CHD) is higher in disadvantaged communities and, as the inverse care law predicts (Hart, 1971; Watt, 2002), more-deprived areas are less likely to have available health services to address CHD. Ashford, Davidson, and Yazbeck (2006) describe several mechanisms that might underlie the inverse care law, including "lack of knowledge, lack of power, inaccessibility of facilities that provide decent care, unresponsive health providers, and the cost of some services" (p. ii). Initiatives such as the primary prevention of HaHP that are focused on redressing health equities need to demonstrate how they respond to the inverse care law.

In 1999, the policy document *Towards a Healthier Scotland* (Scottish Executive) discussed an attack on health inequalities and pointed to the first phase of HaHP as a key means of reducing health inequalities. The first phase of HaHP (between 2000 and 2004) focused on community-level intervention primarily through community partnerships (Blamey et al., 2004), while the second phase (2005 to 2008) developed multiple interventions (e.g., primary prevention, cardiac rehabilitation) that provided services directly to individuals. This article focuses on the second phase of the evaluation (the two phases were completed by different evaluation teams) and describes the evaluation of the primary prevention intervention of HaHP in Phase 2.

To close the inequalities gap in CHD and life expectancy, *Delivering for Health* (Scottish Executive, 2005b) called for greater targeting of health improvement action and resources for those living in the most disadvantaged areas in addition to the development and delivery of "anticipatory care" for those at risk in all areas. The document also set a target to reduce premature mortality in the most disadvantaged communities at a rate 15% greater than the national average. Several other policy documents describe the Scottish government's commitment to addressing inequalities in health (Scottish Executive, 2005a, 2007, 2008). The policy rationale central to HaHP's Phase 2 primary prevention was to encourage individuals to recognize their own role as "co-producers of their own health."

Programmatic Background

Table 1 describes the key steps involved in the primary prevention dimension of HaHP Phase 2. Figure 1 describes the key program logic involved in the primary

Table 1. Description of HaHP Primary Prevention (adapted from Sridharan et al., 2008a, and Pawson & Sridharan, 2009)

Reach Primary prevention of HaHP targeted individuals aged 45–60 years old through delivery of a tailored primary prevention system. The Central Data Repository—a centralized database that linked data from general practices, laboratories, and hospital and national data sets—was used to support the delivery of primary prevention and acted as a sampling frame for the intervention. All eligible individuals (i.e., those aged 45–60, without a history of heart disease, living in Paisley, and enrolled with a Paisley GP/primary care physician) were sent a personal invitation through the mail to attend a "heart health check." Mailing was targeted at the most deprived postcode sectors in the first instance. Subsequent postcode sectors were then offered the intervention. Recipients who accepted the offer to attend a heart health check were required to call a number provided in the invitation letter. The telephone calls were handled by the call centre.
Screening All those who accepted this offer were screened by HaHP's nurses. This enabled their risk of developing coronary heart disease in the next 10 years to be calculated, and individuals were informed of their risk status (low, moderate, or high).
Health Coaching All screening participants were then offered an opportunity to meet with a HaHP Health Coach. Health coaching in HaHP Phase 2 was planned as a method to engage the target population by providing one-to-one client-led support and individualized guidance to empower individuals to make positive lifestyle changes aimed at reducing the risk of developing CHD.
Micro-interventions In addition, health coaches had the option of providing individuals information on either HaHP's own micro-interventions or other appropriate community services identified during a local mapping exercise. Health coaches used a web-based system to facilitate health coaching consultations and to capture details of services for health coaches to use in signposting suitable individuals.

Figure 1. A Simplified Program Logic of Primary Prevention Intervention of HaHP

prevention intervention. The details of the primary prevention program have been described in Sridharan et al. (2008a). The key components of the HaHP primary prevention intervention were Reach, Screening, Health Coaching, and a range of Micro-Interventions. Not only was the overall program a "black-box" but, as described later, each of the components needed greater clarification.

A BRIEF PRIMER ON REALIST EVALUATION

The realist approach attempts to understand why programs work. "Realists do not conceive that programmes 'work,' rather it is action of stakeholders that makes them work, and the causal potential of an initiative takes the form of providing reasons and resources to enable program participants to change" (Pawson & Tilley, 1997, p. 215).

Realist evaluation approaches are especially relevant for evaluating complex, multicomponent interventions such as HaHP. Pawson, Greenhalgh, Harvey, and Walshe (2004) highlight seven characteristics of complex interventions, listed in Table 2. Table 2 also describes questions that might emerge in planning an evaluation from a realist lens. One of the implications of a realist view of complex programs is a recognition that program implementers need help to align the complex programming with the long-term goals (such as redressing health inequities).

A good introduction to how to conduct realist evaluations is found in Pawson and Tilley (2004). We briefly describe how we have approached realist evaluations in our work along with providing a flavour of the results.

1. Explicate Context, Mechanism, and Outcomes (CMO) Configurations

The starting point is clarifying how the program is likely to work: rather than focus on a "boxes and arrows" view that is normally provided by logic models, realist evaluation explores and severely interrogates the contexts and mechanisms that are necessary for the program to work. Key questions that a realist approach to evaluation raises are What is it about a program (the "program mechanisms") that brings about change? and What contexts are needed for program mechanisms to be activated? This step is usually done by working closely (interviews, focus groups, etc.) with multiple program stakeholders in developing the CMO configurations, and scouring official documents and literature reviews.

2. Prioritize Key Links in the Theory of Change

It is unlikely that all of the linkages or all of the CMO configurations can be investigated in a single evaluation. A key step is to work closely with program stakeholders to decide which of the linkages in the theory of change need to be prioritized as part of the evaluation. As example, in the case of HaHP, exploring the component of reach was an important priority from a policy perspective, given the centrality of reaching the poorest individuals as part of redressing health inequities.

Table 2. Features of Complex Interventions (adapted from Sridharan, Dunn, & Nakaima, 2012)

Features of complex interventions (Pawson et al., 2004, p. iv)	**Examples of evaluation questions for health inequities**
"The intervention is a theory of theories."	What are the stakeholders' theories of the intervention? Do different stakeholders have different theories of how the intervention will impact health inequities?
"The intervention involves the actions of people."	How do key stakeholders co-construct the intervention? What are the active ingredients of each of the interventions? Is the actual "journey" of the intervention different from the planned journey? Is there buy-in from the stakeholders for the theory of the intervention? What role do different actors play in the success of a program?
"The intervention consists of a chain of steps."	What are the implications of a complex chain of program activities for impacting long-term outcomes such as health inequities? How do upstream and downstream interventions connect with the causal chain implicit in the intervention? Which step in the chain is especially critical in the impact of programs?
"These chains of steps or processes are often not linear, and involve negotiation and feedback at each stage."	How does user involvement change the planned intervention over time? What role do key users play in the success of the program?
"Interventions are embedded in social systems, and how they work is shaped by this context."	How did the context of the intervention influence the planning and implementation of the intervention? What role did the organizational context play in shaping the intervention?
"Interventions are leaky and prone to be borrowed."	How and why did the intervention change over time? Did the program theory change over time?
"Interventions are open systems and change through learning as stakeholders come to understand them."	How did the experience of implementing a complex intervention change program staff's perceptions of the mechanisms involved in impacting long-term outcomes? What are the implications of such learning for future interventions?

3. Connect the Various CMO Configurations with a Framework of Learning

As discussed below, exploring questions about the configurations of contexts, mechanisms, and outcomes can result in a broad range of "learnings" that includes

knowledge development, policy, organizational, process, and impact learning (Sridharan et al., 2008a). Examples of learnings from the HaHP primary prevention evaluation include

Policy learning: Policy makers typically want to know more than just whether a program works (Sanderson, 2003). For example, in the evaluation of a program such as HaHP it is important to understand what has been learned about the underlying policy assumption that an emphasis on area-level deprivation is the way to address health inequities. Based on the evaluation, how can future versions of similar interventions better translate the underlying policy theory into more clearly developed programs?

Organizational learning: What has been learned about the organizational systems needed to implement a complex program? Note that programs are "complex systems thrust amidst complex systems" (Pawson et al., 2004)—what kinds of organizational structures are needed to successfully implement complex programs such as HaHP? Was coordination planned across the different teams implementing the multiple interventions?

Process learning: Questions about mechanisms also provide an opportunity to learn more about program processes. In HaHP, some examples of key processes that needed greater investigation were how to raise expectations of disadvantaged individuals who already feel marginalized from the "system," which results of the screening should be shared with clients, and how health coaching works.

Knowledge development: In our experience, programs often don't have sufficient knowledge about the lives of the program recipients they are trying to impact. Evaluations provide an opportunity to build such knowledge. For example, in the case of HaHP, should health coaching sessions be set up during the day or in the evening? Knowing the employment status of individuals and their daily routines can be critical to the success of the intervention. The evaluation provides opportunities to learn about the barriers to change; such learnings are often assumed at the start of the program. Such knowledge can be fed into future versions of the program.

Program impacts: The focus of realist evaluation on contexts, mechanisms, and outcomes does not preclude a focus on impacts. Policy-makers want to know the impacts of programs like HaHP on both individuals and communities. Does the intervention work? Are there demonstrable impacts of the interventions on program recipients? Is there empirical support for program impacts on interim and proximal individual-level outcomes based on the program theory? Typically a realist evaluation approach can also be integrated with experimental or quasi-experimental designs in assessing the impacts of programs.

4. Implement a Mixed Methods Design and Collect a Range of Data

A key step is to collect a range of data to explore and interrogate the CMO configurations. Pawson and Tilley (2004, p. 11) provide a detailed account of this step:

> Data gathering has the task of trying to match information to these various leads. Given the preliminary theories cover mechanisms and contexts and outcomes, data collection has to be both qualitative and quantitative. The evaluator has, quite literally, to scavenge for the best data to test out the theories. Existing administrative records might be put to use, stakeholders of all type might be interviewed and shadowed, dedicated before-and-after measures might be designed and put in place, focus groups might be assembled to unearth reasons for crucial choices, and so on.

As an example, our evaluation of "reach" was investigated through multiple data sources including understanding the factors that predicted who was responding to the invitation to join the program, exploring the impact of both individual and area-level deprivation in participating in the intervention (Sridharan et al., 2008a), interviews and focus groups with staff on whether the program was doing enough to address health inequities, and interviews with program recipients on whether the program needed to be modified to address their needs. Table 3 describes the design and methods that supported the evaluation of the primary prevention intervention of HaHP.

5. Analyze, Learn and Revise

Consistent with the mixed data collection/mixed methods design, a range of analytical techniques need to be conducted to address the various learning questions discussed above. Again consider Pawson and Tilley (2004, p. 11) for details of this step:

> Have the theories about how the programme worked been supported or refuted by the proceeding analysis? Judgment on this score is invariably mixed, with some output and outcome variations being clear and intelligible, whilst others remain quite puzzling. Just

Table 3. Designs and Data for HaHP interventions (Sridharan et al., 2008a)

Design	Data
A longitudinal design (without a comparison group); data were collected at multiple time points from both program recipients and program staff; methods implemented were both qualitative and quantitative	Detailed data on program planning and implementation processes were obtained by means of longitudinal interviews and focus groups with HaHP program staff. Detailed data were also collected from program recipients using multiple sources. Data were integrated from a variety of sources to follow individuals over the multiple stages of the intervention. Data sources included • Structured measures on participant behaviours and experiences from the Health Behaviour Change Network • Structured telephone interviews with a large sample of clients • Detailed personal interviews of a small sample of clients • Detailed information on risk factors and risk of CHD for all program participants who attended both the baseline and 12-month follow-up was obtained from the Central Data Repository

as with programme building itself, quite unanticipated effects can be uncovered in the sub-group analysis and these require a revisit to the hypothesis drawing board.

Table 4 describes some examples of learnings that emerged from the evaluation.

Table 4. Some Examples of Learnings from the Evaluation of the Primary Prevention Intervention of HaHP (Sridharan et al., 2008b)

Learning(s)	Examples of learning from the evaluation of the HaHP primary prevention intervention (excerpts from the final report)
Policy learning	"There is an unfortunate tendency in assessing progress of health interventions to focus on the number of individuals that get screened and treated without focusing on the characteristics of these individuals. Such a tendency is encouraged by an uncritical focus on meeting target outputs. Addressing the inverse care law clearly requires that programs address the needs of those who are most in need. An important planning tool for future interventions should be to identify which individuals need to be treated most, which can be done through the development of inequality metrics that gives weighting to the degree of need of participants. A system of incentives needs to be devised for monitoring intervention progress by which delivery, treatment, and engagement with a small group of individuals who need such interventions is rewarded compared to another intervention that treats a considerably larger number of the 'worried well.'" (p. 25)
Organizational learning	"A future program theory also needs to consider the consequences of the time delays between different parts of the intervention chain. As an example, the time delay between screening and health coaching might serve as a barrier to engagement. There is a need to engage people when the opportunity arises (Watt and Sridharan, 2008), and to take advantage of periods of increased motivation." (p. 16)
Process learning	"Health coaches felt that the recruitment strategy used had largely failed to engage the target 'hard to reach' population and had resulted in an overrepresentation of the 'worried well.' Clients often reported being aware of the need to make changes and being motivated to take part because of that existing awareness. The 'hard to reach' population is by definition not ready to respond, and therefore this recruitment is unlikely to be successful for them. For the small proportion of more deprived clients with more complex needs, the fruit and vegetable vouchers were valued. However, staff felt that more could be done to offset costs relating to healthy eating and taking part in exercise, and indeed clients also reported such costs as barriers." (p. 13)

(*Continued*)

Table 4. (Continued)

Learning(s)	Examples of learning from the evaluation of the HaHP primary prevention intervention (excerpts from the final report)
Knowledge development	"HaHP staff raised a number of questions regarding what should be the right level of flexibility in anticipatory care initiatives. While staff clearly valued the principles of 'client-choice' and 'tailoring to individual needs,' some noted that the practical implications of adopting such an ethos need to be fully considered and in particular that the degree of flexibility that is offered by such initiatives should be based on capacity. Terms like 'client-led' need to be fully conceptualized. For example, the extent to which a program is tailored to individual need has implications for engagement, subsequent and long-term outcomes." (p. 16)
Program impacts	"There was evidence of a favourable impact of the primary prevention on program participants for a number of health measures, including risks of CHD. The impact is especially pronounced when the risk associated with age was removed from the Framingham risk scores. The size of the effect suggests that there may have been a 'real effect' in decreasing participants' risk scores. Significantly the reduction in risks of CHD was observed for individuals with high levels of baseline Framingham Risk Score. Also, program participants self-reported many positive lifestyle changes and indicated a greater awareness of the importance of lifestyle changes and healthy behaviours. Both of these factors indicate that the intervention was successful in aiding participants to make changes that decreased their risk of CVD." (p. 11)

Further, working closely with policy makers and program staff provides the opportunity to revise and refine the program to incorporate learnings into future versions of the program.

HOW EVALUATIONS CAN HELP WITH MEETING POLICY GOALS

A key realist insight that informs our thinking is that the policy implementation process is typically a "long-chained" process—consider Pawson et al. (2004, p. 5):

> Intervention theories have a long journey. They begin in the heads of policy architects, pass into the hands of practitioners and managers, and (sometimes) into the hearts and minds of clients and patients. Depending on the initiative, different groups will be crucial to implementation; sometimes the flow from management to staff (and through its different levels) will be the vital link; at other times the participation of the "general public" will be the key interchange.

In this section we discuss the multiple roles that evaluation needs to play along such a long chain. Using the evaluation of HaHP as one example, Figure 2

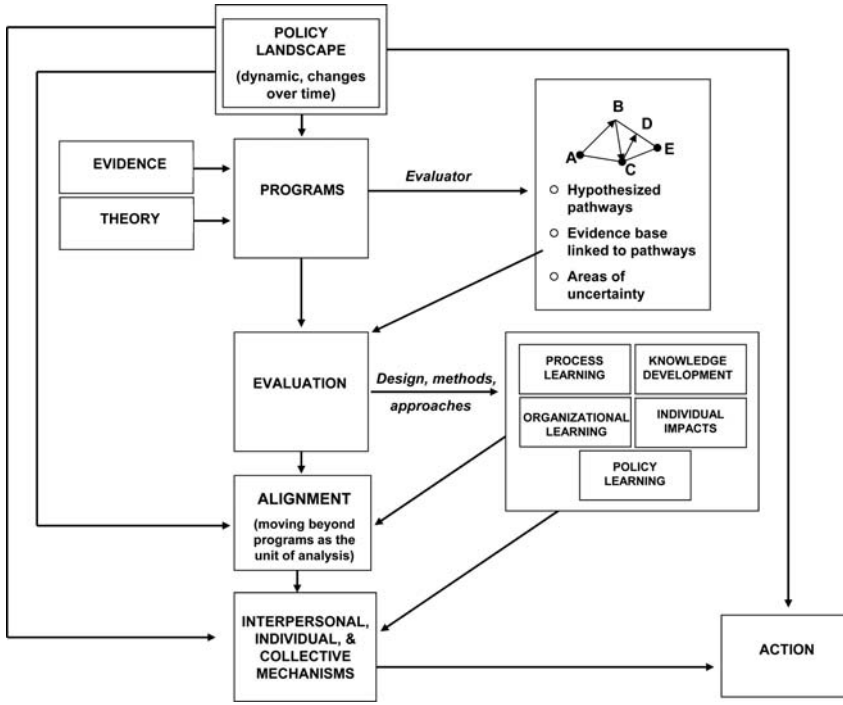

Figure 2. How Realist Evaluation Can Help with Meeting Policy Goals

describes our understanding of the process by which evaluative thinking and evaluative methods can help with policy implementation. Our interest here is not to faithfully model the policy influence process (Carden, 2009) but to identify the different kinds of roles that evaluations can play in influencing the policy implementation process. Figure 2 represents a gross simplification of the policy influence process: a heterogeneous, chaotic, nonlinear process is presented as a linear process of well-defined discrete steps. We recognize that the "reality" of policy influence is considerably more complex, but Figure 2 serves as a useful device to map out the different roles that learnings from evaluations can play in policy influence.

Key roles that evaluation can play include exploring "loss in translation" from policy aspirations to program design; interrogating the program design; developing a range of learnings from conducting the evaluation; and aligning the learning with policy priorities.

1. Explore "Loss in Translation" from Policy Aspirations to Program Design

Translating the policy aspiration into a demonstration project can be a challenge. Thinking about contexts, mechanisms, and outcomes in the focused manner that

realist evaluation promotes can help raise questions about the match between the program design and the longer-term policy aspiration. As described by Sanderson (2003), functions of evaluations can include "influencing the conceptualization of issues, the range of options considered and challenging taken-for-granted assumptions about appropriate goals and activities" (p. 333).

This journey from learning lessons from an evaluation of a demonstration project to impacting health inequities through spread and scale-up is both uncertain and convoluted. Realist evaluation methods with their focus on contexts and mechanisms can help with this journey by raising questions on the implied assumptions of spread and scale-up that both challenge and aid reflection at the early stages of policy implementation—however, this reflection implies that thinking about evaluation happens early in the policy planning/implementation process stage itself. In our experience, evaluators are often brought into the evaluation process late in the game.

2. Interrogating the Program Design

Evaluative thinking around CMO configurations can also help clarify the role of evidence and program theory in designing the program. The focus needs to be on the types of evidence that are useful given the heterogeneous nature of real-world complex systems. For example, Rodrik (2009) makes this important point: "The 'hard evidence' from the randomized evaluation has to be supplemented with lots of soft evidence before it becomes usable" (p. 5). Questions to explore at this stage include, for example, How strong is the evidence base that informed the design of the program? What were the key gaps and uncertainties in the evidence base? Was the evidentiary support for effectiveness in multiple contexts and the choice of mechanisms that informed the program design strong? In our experience, despite the increased attention to evidence-based programming, programs often struggle at the point of implementation. Given that programs are often complex with multiple components and linkages, some aspects of the program have a strong evidence base, while others may not be as strongly evidence-based. In our experience, in all cases, upfront reflection on contexts, mechanisms, and outcomes can help with the program design.

3. Learnings from Conducting the Evaluation

As described earlier, varieties of learning may be possible from this application of a realist evaluation approach. Shifting the focus from "What works?" to "What is it about this program that makes it work?" can assist in developing multiple types of learning. Realist evaluation can fruitfully combine multiple methods, including both qualitative and quantitative approaches (Pawson & Tilley, 2004) as it seeks to understand the relationships between contexts, mechanisms, and outcomes.

4. Aligning the Learning with Policy Priorities

The challenge of "making the evaluation matter" is not just to generate learning(s) but to align such learning with policy priorities. Although there is a growing

literature on external validity (Deaton, 2009; Rodrik, 2009), how learning from evaluations can be spread to impact populations has a limited evidence base. Questions of alignment between the evaluation and policy priorities are especially important given dynamic policy landscapes, changes in government, and changes in policy priorities. Impacting population health is not simply a matter of replicating a proven effective intervention all across a country like Scotland. Populations have heterogeneous needs and part of the complexity in taking programs to scale at the population level requires addressing the diverse needs of the population. A policy framework on inequities needs to go considerably beyond the details of a specific program (Sridharan et al., 2008a). A program theory usually informs why a program can be effective; however, it rarely describes why and how a program can impact inequities at the population level. More formal, explicit, and pragmatic thinking is needed on how and why an intervention such as HaHP can impact health inequities at the population health level. In our view, questions about contexts and mechanisms should not be confined to the program level; raising questions about contexts, mechanisms, and their relationship to equity outcomes at the population level can help enhance policy implementation.

The results of the evaluation by themselves might often not be useful without explicit consideration of mechanisms of policy influence (Carden, 2009; Mark & Henry, 2004). From our experience, key policy initiatives to consider for an evaluation to be influential include

- Fit with the latest "political zeitgeist."

  For example, the political zeitgeist was conducive to implementing health equity policy. In our view, implementing HaHP-type programs might be easier in Scotland than in some other countries, as the view of health inequities as something that a decent cohesive society should seek to reduce is highly valued across the various political parties that have and are likely to form governments. Another aspect of such a zeitgeist was the widely held belief that it was the role of the health services to target the most vulnerable, especially by using the well-developed UK primary care system of general practice, to reduce the worst impacts of health inequities. Such contexts may not be true in other settings. It is helpful for the evaluator to understand the policy landscape (see Figure 2). Working closely with policy makers and other policy and program stakeholders provides one avenue to understand the policy landscape and, in some cases, impact the policy landscape through broadening of policy horizons and enhancing policy capacities (Carden, 2009).

- Raise the salience (Mark & Henry, 2004) of key policy considerations to those with power to decide.
- Be simple enough to be communicable.
- Be simple enough for those delivering to be able to apply it. Realist evaluation embodies the principle of "nothing as practical as a good theory" (Pawson, 2003).

- Be cogent enough to persuade the various agents in the delivery chain that it is worth doing.
- Be cogent enough to win over the required leadership—from the start, from the top, to the end.
- Be embedded within intervention teams. For example, explore perspectives with action research approaches, which are helpful but not captured by the intervention team—with people who can remind the intervention team about the gap between their measures and the intended outcome. Such embedded evaluators can mediate or make present the experience of others in similar contexts, and advise on what is more or less likely to work.
- Educate policy makers and practitioners that evaluations have very different functions and are not simply instruments to determine if a program is working.

CONCLUSION

Key strengths of a realist approach include a framework of learning that promotes diverse types of learning, which will appeal to a wide range of stakeholders; an explicit focus on the complexity of the implementation chain; and a framework that seeks to understand why programs succeed or fail.

An approach like realist evaluation is likely to lead toward positive impacts if

(a) ***The evaluator is involved early in the planning and implementation stage.*** There needs to be recognition of the long chains involved in policy implementation, and evaluative thinking is needed at different parts of the implementation. Achieving success across the different links in the implementation chain will require multiple mechanisms to be planned, coordinated, and implemented across the multiple contexts. This implies that ideally the evaluation is initiated right while the program is being planned. We don't think realist evaluation is the right evaluation approach if the evaluation is considered and commissioned after the implementation of a program.

(b) ***Dialogue spaces for ongoing learning are explicitly planned***. While a constructive attitude toward positive impact is helpful, we think there need to be spaces in which co-learning between the stakeholders is possible. Dialogue spaces need to be created for interaction between the multiple stakeholders. Thinking about contexts and mechanisms typically requires dynamic interaction between program planners, implementers, and evaluators. This means moving away from a purely product view of knowledge ("When can we have the final report?") and encouraging spaces in which dialogue and ongoing learning is encouraged. Learning needs to be supported by both well-thought-out processes and products. Learning processes need active and intentionally planned spaces for ongoing dialogue.

(c) ***There is an interest in pluralistic methods and designs***. One of the strengths of realist evaluation is that its focus on context, mechanisms, and out-

comes creates a platform for multiple methods and is not necessarily exclusively qualitative or quantitative. In our experience, evaluators spend more time quibbling about the best methods and, surprisingly, somewhat less time and energy thinking about influence. In our experience, such methodological territorial wars might interfere with the process of policy influence. Realist evaluation with its focus on "what works, for whom, and under what contexts" can promote a pluralistic approach to learning.

REFERENCES

Ashford, L. S., Davidson, R. G., & Yazbeck, A. S. (2006). *Designing health and population programs to reach the poor*. Population Reference Bureau. Retrieved from http://www.prb.org/pdf06/DesigningPrograms.pdf

Blamey, A., Ayana, M., Lawson, L., Mackinnon, J., Paterson, I., & Judge, K. (2004). *Final report of the independent evaluation of the National Health Demonstration Project Have a Heart Paisley (Phase One)*. Public Health and Health Policy, University of Glasgow. Retrieved from www.scotland.gov.uk/Publications/2005/03/20836/54354

Carden, F. (2009). *Knowledge to policy: Making the most of development research*. New Delhi, India: Sage.

Deaton, A. (2008; revised 2009). *Instruments of development: Randomization in the tropics, and the search for the elusive keys to economic development*. The Keynes Lecture, British Academy.

Mark, M. M., & Henry, G. T. (2004). The mechanisms and outcomes of evaluation influence. *Evaluation*, *10*(1), 35–57. http://dx.doi.org/10.1177/1356389004042326

Pawson, R. (2003). Nothing as practical as a good theory. *Evaluation*, *9*(4), 471–490. http://dx.doi.org/10.1177/1356389003094007

Pawson, R., Greenhalgh, T., Harvey, G., & Walshe, K. (2004). *Realist synthesis: An introduction* (Paper 2). ESRC Research Methods Programme, University of Manchester. Retrieved from http://betterevaluation.org/sites/default/files/RMPmethods2.pdf

Pawson, R., & Sridharan, S. (2009). Theory-driven evaluation of public health programmes. In A. Killoran & M. Kelly (Eds.), *Evidence-based public health: Effectiveness and efficiency* (pp. 43–61). Oxford, UK: Oxford University Press.

Pawson, R., & Tilley, N. (1997). An introduction to scientific realist evaluations. In E. Chelimsky & W. R. Shadish (Eds.), *Evaluation for the 21st century: A handbook* (pp. 405–418). Thousand Oaks, CA: Sage.

Pawson, R., & Tilley, N. (2004). *Realist evaluation, magenta book*. Strategy Unit Cabinet Office. Retrieved from http://www.communitymatters.com.au/RE_chapter.pdf

Rodrik, D. (2009). The new development economics: We shall experiment, but how shall we learn? In J. Cohen & W. Easterly (Eds.), *What works in development? Thinking big and thinking small* (pp. 24–47). Washington, DC: Brookings Institution Press.

Sanderson, I. (2003). Is it "what works" that matters? Evaluation and evidence-based policy making. *Research Papers in Education*, *18*(4), 329–343.

Scottish Executive. (1999). *Towards a healthier Scotland: A white paper on health*. Edinburgh, UK: Author.

Scottish Executive. (2005a). *Building a health service fit for the future*. Retrieved from http://www.scotland.gov.uk/Resource/Doc/924/0012113.pdf.

Scottish Executive. (2005b). *Delivering for health*. Retrieved from http://www.scotland.gov.uk/Resource/Doc/76169/0018996.pdf

Scottish Government. (2007). *Better health, better care: Action plan*. Retrieved from http://www.scotland.gov.uk/Resource/Doc/206458/0054871.pdf

Scottish Government. (2008). *Equally well: Report of the Ministerial Task Force on Health Inequalities*. Retrieved from http://www.scotland.gov.uk/Publications/2008/06/25104032/16.

Sridharan, S., Dunn, J. R., & Nakaima, A. (2012). Addressing health equities in social epidemiology: Learning from evaluation(s). In P. O'Campo & J. R. Dunn (Eds.), *Rethinking social epidemiology: Towards a science of change* (pp. 247–263). Springer.

Sridharan, S., Gnich, W., Moffat, V., Bolton, J., Harkins, C., Hume, M., . . ., & Docherty, P. (2008a). *Evaluation of primary prevention intervention: Have a Heart Paisley Phase 2*. Glasgow, UK: NHS Health Scotland.

Sridharan, S., Gnich, W., Moffat, V., Bolton, J., Harkins, C., Hume, M., . . ., & Docherty, P. (2008b). *Learnings from the independent evaluation of Have a Heart Paisley Phase 2*. Glasgow, UK: NHS Health Scotland.

Tudor Hart, J. T. (1971). The inverse care law. *Lancet*, *297*(7696), 405–412. http://dx.doi.org/10.1016/S0140-6736(71)92410-X

Watt, G. (2002). The inverse care law today. *Lancet*, *360*(9328), 252–254. http://dx.doi.org/10.1016/S0140-6736(02)09466-7

Watt, G., & Sridharan, S. (2008). *Towards a definition of anticipatory care*. Edinburgh, UK: University of Edinburgh, Evaluation Programme, Research Unit in Health, Behaviour and Change.

AUTHOR INFORMATION

Sanjeev Sridharan is Director of the Evaluation Centre at the Li Ka Shing Knowledge Institute at St. Michael's Hospital and Associate Professor in the Department of Health Policy, Management and Evaluation at the University of Toronto. Prior to his appointment at Toronto, he was the Head of the Evaluation Programme and Senior Research Fellow at the Medical School of the University of Edinburgh. He is on the Board of the *Canadian Journal of Program Evaluation* and the *Journal of Evaluation and Program Planning*, and also served as Associate Editor of the *American Journal of Evaluation*.

Tim Warren went the University of Aberdeen in 1980, qualifying as a social worker in 1986. He worked in residential child care, before returning to education, in New York and then Edinburgh. He worked in criminal justice social work in Aberdeen and later in Dumfries, before becoming a social work inspector in the Scottish Executive in 2000. Here he provided professional advice, moving from offenders' work to the team that established the Scottish Social Services Council and the Care Commission, and then to the reform of social work education. He cowrote guidance on restraining children in

residential and secure care, and then moved to a policy role, dealing with child protection training. In 2007 he took on the role of policy lead for health improvement and health inequalities, with responsibility for the Keep Well program of inequalities-targeted high-risk primary prevention health checks. At the beginning of 2011 he moved to the Healthcare Planning Division, where he is responsible for the Scottish Government's strategic partnership with the Long Term Conditions Alliance Scotland and for promoting Self Management.

doi: 10.3138/cjpe.29.2.87
